ACADEMIC PRESS

Protein Expression and Purification 25 (2002) 270–282

Protein Expression & Purification

www.academicpress.com

# Gene optimization is necessary to express a bivalent anti-human anti-T cell immunotoxin in *Pichia pastoris*

Jung Hee Woo,[a] Yuan Yi Liu,[a] Askale Mathias,[a] Scott Stavrou,[a] Zhirui Wang,[a] Jerry Thompson,[b] and David M. Neville, Jr.*

[a] *Section on Biophysical Chemistry, Laboratory of Molecular Biology, National Institute of Mental Health, 36 RM 1B08, 36 Convent Drive, Bethesda, MD 20892-4034, USA*

[b] *Fenske Laboratory, University Park, PA 16802, USA*

Received 18 September 2001, and in revised form 1 February 2002

**Abstract**

The bivalent anti-human anti-T cell immunotoxin A-dmDT390-bisFv($G_4S$) was developed for treatment of T cell leukemia, autoimmune diseases, and tolerance induction for transplantation. The multi-domain structure of the bivalent immunotoxin hinders efficient production in *Escherichia coli* and most eukaryotes are sensitive to the toxin. However, *Pichia pastoris* has a tolerance to levels of DT (diphtheria toxin) that were previously observed to intoxicate wild type eukaryotic cells, including *Saccharomyces cerevisiae*. This tolerance has permitted the optimization of the secreted expression of A-dmDT390-bisFv($G_4S$) in *P. pastoris* under the control of *AOX1* (alcohol oxidase 1) promoter. The original DNA sequence of A-dmDT390-bisFv($G_4S$) was not expressed in *P. pastoris* because of several AT-rich regions, which induce an early termination of transcription. After DNA rebuilding for abolishing AT-rich regions and codon optimization, the immunotoxin could be expressed up to 10 mg/L in the shake flask culture. No differences in the expression levels of immunotoxin were observed by using different secretional signal sequences, $Mut^s$ (methanol utilization slow phenotype) or $Mut^+$ (methanol utilization plus phenotype) phenotypes. Buffered complex medium (pH 7.0) having 1% casamino acids provided the highest expression in shake flask culture and PMSF (phenylmethylsulfonyl fluoride) in the range of 1 to 3 mM further improved the expression level presumably by inhibiting protein degradation. The immunotoxin was purified by DEAE (diethylaminoethyl) Sepharose ion exchange chromatography and Protein L affinity chromatography. The immunotoxin purified from *P. pastoris* culture was as fully functional as that expressed in a toxin resistant mutant CHO (Chinese hamster ovary) cell line. Our results demonstrate that *P. pastoris* is an ideal system for expression of toxin-based fusion proteins. 

Anti-T cell immunotoxins were developed for treatment of T cell leukemia, autoimmune diseases, and tolerance induction for transplantation. Recent studies show that a short course of anti-T cell immunotoxin combined with deoxyspergualin results in long-term tolerance of monkey kidney and pancreatic islet allografts with freedom from chronic rejection [1,2]. These immunotoxins were produced by optimized chemical conjugation of a diphtheria toxin (DT)[1]-binding mutant (CRM9) and an anti-CD3 antibody and can kill 2–3 logs of T cells in vivo [3,4]. Because these chemical conjugates are produced in low yields and are heterogeneous in the placement of the toxin and antibody moieties, it is advantageous to replace them with recombinant immunotoxins. Recombinant immunotoxins produced in our lab retain the first 390 amino acid residues of DT consisting of the catalytic A chain and the translocation domain of the B chain fused to a C terminal antigen-binding domain (sFv) of anti-CD3ε antibody, UCHT1. CD3ε is a signaling protein within the T cell receptor

* Corresponding author. Fax: +1-301-402-0245.
*E-mail address:* davidn@helix.nih.gov (D.M. Neville, Jr.).

[1] *Abbreviations used*: A-dmDT390-bisFv($G_4S$), bivalent anti-human T-cell immunotoxin; DT, diphtheria toxin; AOX1, alcohol oxidase I gene of *Pichia pastoris*; PMSF, phenylmethylsulfonyl fluoride; CHO, Chinese hamster ovary cell line; EF-2, elongation factor 2; sFv, single chain Fv fragment; CRM9, diphtheria toxin binding mutant; DT390, amino terminal 390 amino acid residue of diphtheria toxin; dmDT390, DT390 having double mutations on the two potential *N*-glycosylation sites; $(G_4S)_3$ or $(G_4S)$, flexible linker containing GGGGSGGG GSGGGGS; $Mut^S$, methanol utilization slow phenotype of *P. pastoris*; $Mut^+$, methanol utilization plus phenotype (wild type) of *P. pastoris*.

1046-5928/02/$ - see front matter 
PII: S1046-5928(02)00009-8

complex [5,6]. The DT390 truncation has been demonstrated to be the minimum size of truncated DT present in the fusion immunotoxin needed to intoxicate T cells and halt protein synthesis (Thompson and Neville, unpublished data). This immunotoxin was originally produced in *E. coli* utilizing the *Corynebacterium diphtheriae* signal sequence. In the absence of refolding, the potency of the recombinant immunotoxin was 6.25% of the chemical conjugate [6]. The potency could be increased 16-fold by a standard refolding protocol [5,7] or by expressing the immunotoxin in CHO cells that had been mutated to toxin resistance [5,8]. Wild type CHO cells transfected with an immunotoxin gene grew poorly and failed to express immunotoxin [8]. (All eukaryotes are susceptible to protein synthesis inhibition by DT if the toxin can be transported to the cytosol compartment where the toxin substrate elongation factor 2 (EF-2) is ADP-ribosylated and inactivated [9].) CHO cell expression of active immunotoxin required the removal of two potential toxin *N*-glycosylation sites [8]. Although the potency of this recombinant immunotoxin was equal to that of the chemical conjugate, binding studies to T cells showed that the binding affinity of the DT390-sFv construct was 3.3% of the chemical conjugate and 1% of the parental antibody [5]. This appears to be caused by the steric effect of a large protein domain $NH_2$-terminal to the sFv and may be unique to binding to the CD3ε epitope. We have made various versions of recombinant anti-T cell immunotoxin for improving the binding and toxicity [5]. Among these constructs, the bivalent single chain immunotoxin A-dmDT390-bisFv ($G_4S$) construct showed the highest binding and toxicity. It is constructed with the first 390 amino acids of DT (DT390) from the $NH_2$-terminus and two sFv domains that are linked by a $(G_4S)_3$ linker. In this construct, two mutations were introduced for removing the potential *N*-glycosylation sites at positions 16–18 in the DT A chain and positions 235–237 in the translocation domain of the DT B chain. The notation for this double mutation is the letters dm preceding DT390. Glycosylation at these two positions markedly reduced toxicity that was restored by the mutations [8]. Since this bivalent immunotoxin had a multi-domain structure with 5 disulfide bonds, it proved difficult to purify a bioactive fusion protein by in vitro refolding in the *E. coli* system. Therefore the A-dmDT390-bisFv ($G_4S$) was produced in the toxin resistant CHO cell line K1 RE1.22c [5].

In the course of attempting to optimize A-dmDT390-bisFv ($G_4S$) production in toxin resistant CHO cells, it became apparent that 24 and 48 h media concentrations of immunotoxin were limited to a 2–5 μg/ml ceiling by a feedback inhibition mechanism (Mathias and Neville, unpublished data). This was documented by adding either A-dmDT390-sFv or A-dmDT390-bisFv ($G_4S$) to cells secreting either A-dmDT390-sFv or A-dmDT390-bisFv ($G_4S$) and observing the concentration of the immunotoxins by calibrated SDS non-reducing gels. The difference in mobility of the A-dmDT390-sFv versus the A-dmDT390-bisFv($G_4S$) allowed one to clearly distinguish between added immunotoxin and secreted immunotoxin when the added construct differed from the secreted construct. Both constructs were inhibitory in the secretion tests at 5 μg/ml. This may be due to the toxin moiety that is capable of inserting into membranes and initiating pore formation at pH 6.0 and below [10]. This could occur by retrograde transport of immunotoxin into the secretory pathway with resulting damage to the secretory process. Since we were aware that intact yeast cells (*Saccharomyces cerevisiae*) are resistant to high concentrations of diphtheria toxin added to medium [11], we reasoned that yeast might not be susceptible to this type of retrograde feedback inhibition. We therefore investigated the use of yeast as an expression system for DT based immunotoxins and for this we chose *Pichia pastoris*.

Many studies (reviewed in [12]) have demonstrated that heterologous proteins derived from different origins were successfully expressed as a secreted form in the *P. pastoris* expression system. For *P. pastoris*, protein processing, folding, and post-translational modifications take place successfully inside the cell, leading to a functional protein, which can be secreted into the medium. *P. pastoris* is easy to manipulate and has a lower production cost and a higher expression level than other eukaryotic expression systems. More importantly, we determined that as is the case for *S. cerevisiae*, DT and immunotoxin are not toxic to *P. pastoris* if they are present outside the cell. However, AT-rich regions in a native gene may cause early termination of transcription [13,14]. In this case, DNA rebuilding for abolishing AT-rich regions is required for expression of the gene of interest. Recent reports [14–18] have shown that codon optimization can improve expression levels in various expression systems. Codon optimization was considered to improve expression by using DNA rebuilding for replacement of rare codons to preferred codons for efficient translation in *P. pastoris*.

In this study, we demonstrate that *P. pastoris* can express functional immunotoxins without any mutation to confer a resistance to DT, and that gene % AT reduction is required for efficient expression. This paper also reports optimization of media composition and purification of recombinant bivalent immunotoxin.

## Materials and methods

### *Strains and culture media*

The *E. coli* Nova Blue strain (Novagen) was used as the host for DNA manipulations and grown in low salt LB medium (1% tryptone, 0.5% yeast extract, 0.5%

NaCl; 2% agar in plates; 50 μg/ml ampicillin or 25 μg/ml zeocin for the selection of plasmids). The *P. pastoris* GS115 (*his4*), X-33 or KM71 (*his4 arg4 aox1Δ*::*ARG4*) strains (Invitrogen) were used for production of immunotoxins. *P. pastoris* was grown in YPD medium (1% yeast extract, 2% peptone and 2% dextrose; 2% agar in plates; 100 μg/ml zeocin for selection of transformants) for growth. BMMYC medium (1% yeast extract, 2% peptone, 100 mM potassium phosphate, pH 7.0, 1.34% yeast nitrogen base without amino acids, $4 \times 10^{-5}$% biotin, 0.5% methanol and 1% casamino acids) was used as induction medium for production of immunotoxins or DT A chain. For the selection of $His^+$ transformants, histidine-deficient RD medium (1 M sorbitol, 2% dextrose, 1.34% yeast nitrogen base without amino acids, $4 \times 10^{-5}$% biotin, and 0.005% amino acids mixture) was used. RD agar plate, MM agar plate (1.34% yeast nitrogen base without amino acids, $4 \times 10^{-5}$% biotin, 0.5% methanol and 2% agar) or MMH agar plate (MM plus 0.004% histidine) was used for scoring Mut or His phenotype of the transformants.

*DNA rebuilding*

The pET17b vector (Novagen) was used as the cloning vector for DNA rebuilding. DNA rebuilding was done by PCR with primers listed in Table 1. These primers were designed to lower the AT content of DNA sequences in selected regions as well as for codon optimization for efficient translation. Each primer had a new DNA sequence for rebuilding and/or a specific enzyme site for easy subcloning of each PCR product into the parental vector. In cases where more than 30 bp of rebuilding was required at the 5′ and/or 3′ ends, at least 2 PCR were carried out in series. Rebuilt DNA sequences were confirmed by DNA sequencing.

The construction of different versions of synthetic genes coding for immunotoxin is detailed in Figs. 1 and 2. Fig. 1 details the scheme used to rebuild the DNA sequence in selected regions. The gene of interest was divided into 2 regions consisting of DT390 and bisFv that were separately rebuilt. Briefly, the original recombinant plasmid consisting of the A-dmDT390-bisFv($G_4S$) gene, which was an intermediate construct for CHO cell expression, was used as the template DNA for the first PCR. For the first version of the synthetic gene in which the 3′ end region of the DT390 moiety was replaced with a new sequence, 5 PCR were carried out to obtain the DTat1 PCR product by using primer sets as indicated in Table 1 (see 1 in Fig. 1). After each PCR, the product was purified by gel elution and then the purified PCR product was used as the template DNA for the next PCR. For the second version (2 in Fig. 1), the two PCR products, DTat2-1 and DTat2-2 were amplified by PCR using primers for each PCR product and the DTat1 fragment as the template DNA. The two PCR products, DTat2-1 and DTat2-2, were digested with *Kpn*I and *Xba*I and *Avr*II and *Nco*I, respectively, and subcloned into the *Kpn*I and *Nco*I sites of the parental vector, which was the pET17b vector containing the immunotoxin gene DT390-bisFv($G_4S$). For the third version (3 in Fig. 1), the second version of synthetic gene was used as the template DNA to obtain the first PCR product. Five successive PCR were carried out to amplify the DTat3 PCR product. The PCR product was digested with *Xho*I and *Nco*I and subcloned into the same sites of the parental vector. Each version of the synthetic gene was confirmed by DNA sequencing, inserted into *P. pastoris* expression vectors pnPICZα or pPIC9K and transformed into *P. pastoris* by electroporation. The recombinant immunotoxin, A-dmDT390-bisFv($G_4S$), was expressed in *P. pastoris* only after rebuilding the 3 major AT-rich regions; however, the expression level of immunotoxin was still low.

We then focused on rebuilding the bisFv domains. It was observed that there were four AT-rich regions in the bisFv domains and two minor AT rich regions in the DT390 moiety. These results suggested that there was room for improvement in the expression level. DNA rebuilding work produced many intermediate constructs for making the final construct (the fifth version).

For the fourth version, the three PCR products, DTop1, DTop2, and DTop3, were amplified by using the third version of the synthetic gene and a primer set for each PCR product as indicated in Table 1 (see Fig. 2, panel A). These primer sets were designed to reduce the AT content of two minor AT-rich regions as well as optimize the codon bias for *P. pastoris* by using preferred codons. A preferred codon was defined as a codon having more than a 30% frequency for encoding an amino acid in highly expressed proteins in *P. pastoris* based on the data of Sreekrishna [19]. The preferred codons used for this rebuilding are listed in Table 2. The parental vector digested with *Hin*dIII and *Nco*I was ligated with enzyme-treated DTop1 and DTop2 fragments as indicated in Fig. 2 (panel A) to generate an intermediate construct, DTop1-DTop2-bisFv($G_4S$) in pET. The resulting construct was digested with *Avr*II and *Nco*I and then ligated with the DTop3 fragment that was digested with *Xba*I and *Nco*I, resulting in the fourth version of the synthetic gene construct, DTop390-bisFv($G_4S$) in pET.

For DNA rebuilding of the bisFv region the preferred codons in Table 2 were used in the indicated regions (Fig. 2, panels B and C). The third version of the synthetic gene construct was used as the DNA template to obtain the first PCR products for sFvop1, sFvop2, sFvop3, and sFvop4. The intermediate construct A-dmDT390-N-B-E in pET was digested with *Nco*I and *Bam*HI, and then ligated with the sFvop1 PCR fragment that was treated with *Nco*I and *Bam*HI. The resulting construct A-dmDT390-sFvop1-B-E in pET was treated

Table 1
List of DNA sequence of primers for DNA rebuilding

| PCR Product | Name and DNA sequence of primer |
|---|---|
| DTat1 (5 PCR) | 5′*Kpn*-E-DT1: CCGGGTACCGGAATTCGGCGCTGATGATGTTGTTGATTCT<br>3′DTat348: TCACCGACCAATGGGATAGCTTGAGCAACCATCAAAGAGGACAAAGCGATGGATTGTGCC<br>3′DTat362: CTCGACGAAGTTGTAGGCAGCGAAACCGATGTCAACCAACTCACCGACCAATGGG<br>3′DTat372: GTGGACGACTTGGAACAAGTTGATGATGGACTCGACGAAGTTGTAGGCAG<br>3′DTat383: CTGGGGAGTAAGCCGGACGGTTGTAGGAGTTGTGGACGACTTGGAACAAG<br>3′*Nco*I-DTat390: GTCCCATGGCAAGAATGGTTGGGTCTTGTGACCTGGGGAGTAAGCCGGAC |
| DTat2-1 (2 PCR) | 5′*Kpn*-E-DT1: CCGGGTACCGGAATTCGGCGCTGATGATGTTGTTGATTCT<br>3′DTat236: CTTAGCGGGGGATTCGGACATCTTGTTCTTGATTGGGCCATGCTCTTTCA<br>3′*Avr*II-DTat248: CCTAGGTATTGCTTGGCCTTTTCCTCGGAGACGGTCTTAGCGGGGGATTC |
| DTat2-2 (1 PCR) | 5′*Xba*I-DTat246: TCTAGAAGAGTTCCACCAAACCGCCTTGGAGCATCCTGAATTGTCA<br>3′*Nco*I-DTat390: GTCCCATGGCAAGAATGGTTGGGTCTTGTGACCTGGGGAGTAAGCCGGAC |
| DTat3 (5 PCR) | 5′DTat47: ACGACGACGACTGGAAGGGGTTCTACTCCACCGACAACAAGTACGACGCT<br>5′DTat34: TATCCAGAAGCCAAAGTCCGGCACCCAAGGTAACTACGACGACGACTGGA<br>5′DTat23: ACCAAGCCAGGTTACGTCGACTCCATCCAGAAGGGTATCCAGAAGCCAAA<br>5′DTat11: CCTTCGTCATGGAGAACTTCGCTTCCTACCACGGGACCAAGCCAGGTTAC<br>5′Kp-E-DTat1: CGGTACCGGAATTCGGCGCTGATGATGTCGTCGACTCCTCCAAGTCCTTCGTCATGGAGAAC<br>3′*Avr*II-DTat248: CCTAGGTATTGCTTGGCCTTTTCCTCGGAGACGGTCTTAGCGGGGGATTC |
| DTop1 (3 PCR) | 5′*Hin*dIII-X-KR-A-DTop1: GATTCAAGCTTGCTCGAGAAGAGAGCTGGTGCTGACGACGTCGTCGAC<br>3′DTop97: GTCGACCTTCAAAGCCAAGACCTTAGTCAGACCTGGGTAGGTGACCTTGA<br>3′DTop109: GGACAAACCCAACTCCTTCTTGATGGTCTCAGCGTTGTCGACCTTCAAAG<br>3′*Kpn*I-DTop120: CTTGGTACCGACTTGCTCCATCAATGGCTCAGTCAAGGACAAACCCAACT |
| DTop2 (2 PCR) | 5′*Kpn*I-DTop119: GTCGGTACCGAAGAGTTCATCAAGAGATTCGGTGACGGTGCTTCCAGAGT<br>5′DTop130: GGTGCTTCCAGAGTCGTCTTGTCCTTGCCATTCGCTGAGGGTTCTTCTAG<br>3′DTop198: GAGGAACCAACAGAACGTCTGACTCTGTTACCAGCACAGGCTTGA<br>3′*Nco*I-Avr-DTop205: ATGCCATGGCCAGCCTAGGTTGATACAGGACAAAGAGGAACCAACAGAAC |
| DTop3 (1 PCR) | 5′*Xba*I-DTop203: CAATCTAGACTGGGACGTCATCAGAGACAAGACTAAGACCAAGATCGAGT<br>3′*Nco*I-DTat390: GTCCCATGGCAAGAATGGTTGGGTCTTGTGACCTGGGGAGTAAGCCGGAC |
| sFvop1 (2 PCR) | 5′*Nco*I-sFvop1: TTGCCATGGGACATCCAGATGACCCAGACC<br>3′sFvop55: GTGCAATCTGGAAGTGTAGTAGATCAACAACTTGACAGTACCGTCTGGCT<br>3′*Bam*HI-sFvop65: ACCGGATCCAGAGAACTTAGATGGGACACCAGAGTGCAATCTGGAAGTGT |
| sFvop2 (1 PCR) | 5′*Bgl*II-sFvop64: TCTAGATCTGGTTCTGGTACTGACTACTCCTTGACCATCTCCAACTTGGA<br>3′*Bam*HI-sFvop117: TCCGGATCCACCACCTCCAGAACCTCCTCCACCCTTGATCTCCAACTTGG |
| sFvop3 (1 PCR) | 5′*Bgl*II-sFvop116: GGTAGATCTGGAGGTGGAGGTTCTGAGGTTCAATTGCAACAATCTGGACC<br>3′*Xho*I-sFvop181: TAGCTCGAGACACCCTTGTAAGGGTTAATCAAACCCATCCAC |
| sFvop4 (1 PCR) | 5′*Sal*I-sFvop179: TGTGTCGACCTACAACCAGAAGTTCAAGGACAAGGCTACTTTCACTGT<br>3′*Bam*HI-sFvop249: TCCGGATCCTCCACCTCCAGAGGAGACAGTGACAGTAGTACCAGCACCCC |
| Second sFvop (1 PCR) | 5′*Bgl*-L10-sFvop1: GAAGATCTGGAGGAGGTGGTTCTGGTGGTGGAGGTTCTGACATCCAGATG<br>3′E-stop-sFvop503: CCGGAATTCTTAAGAGGAGACAGTGACAGTAGT |

with *Bam*HI and ligated with the sFvop2 PCR fragment that was digested with *Bgl*II and *Bam*HI, producing the construct A-dmDT390-sFvop1-sFvop2-B-E in pET which still had a unique *Bam*HI site for further cloning. The resulting construct was digested with *Bam*HI and then ligated with the sFvop3 and sFvop4 PCR fragments that were treated with *Bgl*II and *Xho*I, and *Sal*I and *Bam*HI, respectively. The resulting construct A-dmDT390-$NH_2$-sFvop-B-E in pET was used as the DNA template and cloning vector for the COOH terminal sFv fragment rebuilding. A-dmDT390-$NH_2$-sFvop-B-E in pET was digested with *Bam*HI and *Eco*RI and ligated with the COOH terminal sFvop fragment. The rebuilt DNA sequence of bisFv was excised with *Nco*I and *Eco*RI and then ligated at the *Nco*I and *Eco*RI sites in the fourth version of the synthetic gene construct, resulting in the fifth and final version of the synthetic gene construct. Each version of the synthetic gene from the second DNA rebuilding was inserted at *Xho*I and *Eco*RI sites in pnPICZα and transformed into *P. pastoris*.

*Gene construction for immunotoxin and A chain of DT*

Prior to subcloning, the pPICZα vector was modified to ease the subcloning procedure and to replace the

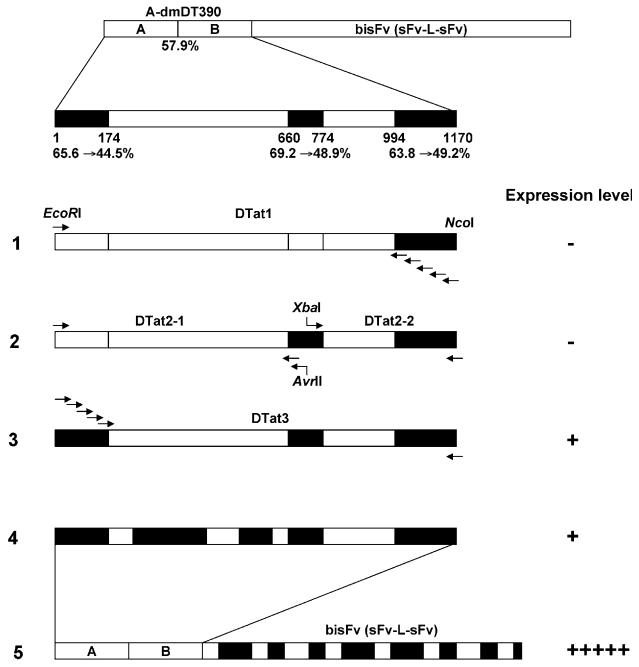


Fig. 1. Different versions of the synthetic gene coding for bivalent anti-human anti-T-cell immunotoxin, A-dmDT390-bisFv($G_4S$). Solid boxes indicate DNA sequences chosen for rebuilding. Percentages represent the AT content of the DNA sequence before and after rebuilding. In the A-dmDT390 region of the first, second, and third versions, DTat1, DTat2-1, DTat2-2, and DTat3 represent the names of PCR products and arrows indicate primer position (see Table 1 for primers).

prepro signal sequence of alpha-mating factor with the signal sequence of killer toxin from *Kluyveromyces lactis* or the pre signal sequence (the amino terminal 19 amino acid residues) of alpha mating factor. Additionally, the pPICZ vector was digested with *Sal*I and *Xho*I, and religated to obtain the pnPICZ vector, which had a deleted *Sal*I–*Xho*I fragment in the multi-cloning site. Simultaneously, the *Hin*dIII–*Eco*RI fragment containing the prepro signal sequence of alpha-mating factor obtained from pPICZα by digesting with *Hin*dIII and *Eco*RI was inserted at the same site of the pnPICZ vector. The resulting vector, pnPICZα, has a unique *Xho*I site upstream of the DNA sequence encoding the Kex2 recognition site (KR). The rebuilt gene of immunotoxin was subcloned into the *Xho*I and *Eco*RI sites of pnPICZα.

For replacement of the prepro signal sequence of alpha mating factor with the signal sequence of killer toxin [20,21], two primers were designed to be annealed at the 3′ end of each primer and then extended by DNA sequenase (US Biochemical) to obtain the synthetic gene encoding signal peptide of killer toxin. The two primers were; kt1, TAA TTA TTC GAA ACG ATG AAT ATA TTT TAC ATA ttt ttg ttt ttg ctg tca ttc gtt ca; kt2, GGA ATT CAG CTT CAG CCT CAC CTt gaa cga atg aca gca aaa aca aaa a; underlined letters and lower cased

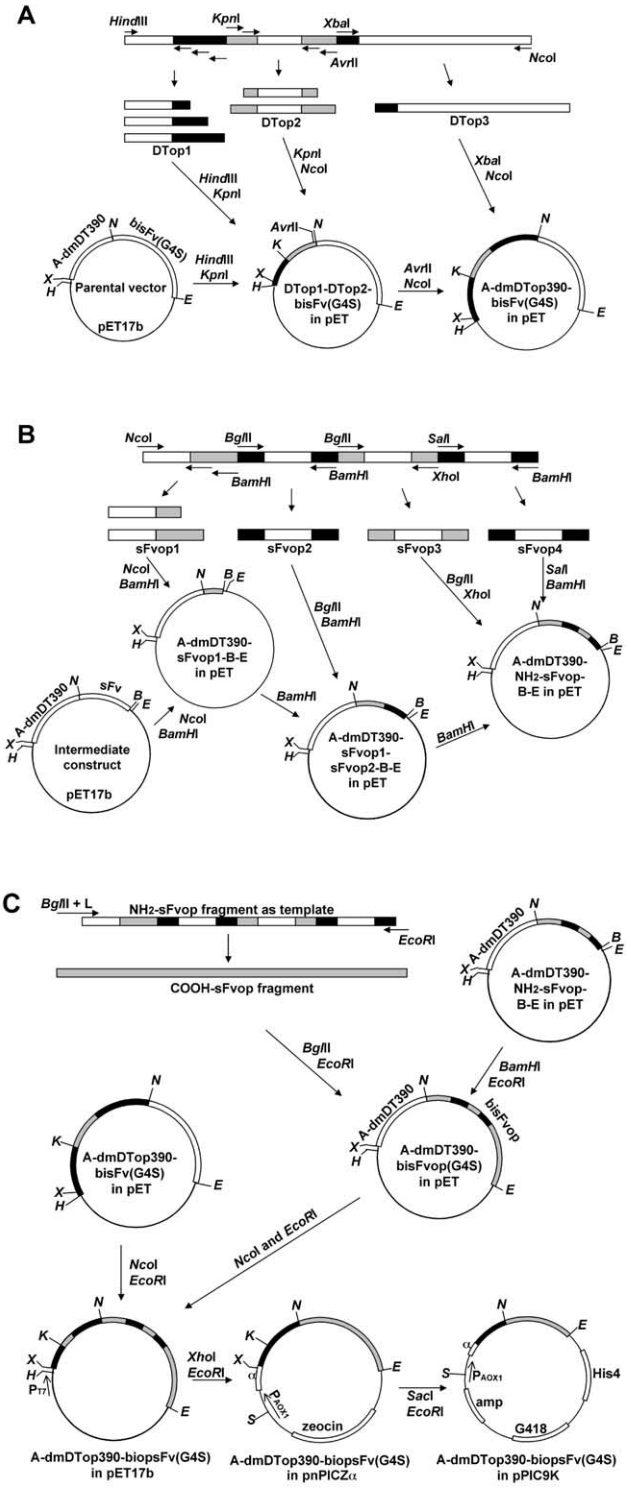


Fig. 2. DNA rebuilding of the bivalent anti-human anti-T-cell immunotoxin gene, A-dmDT390-bisFv($G_4S$). Solid and gray boxes indicate the DNA sequences chosen for rebuilding by different primer sets. A, rebuilding of DT390; B, rebuilding of the amino terminal sFv fragment; C, rebuilding of the carboxyl terminal sFv fragment and subcloning into pnPICZα and pPIC9K vectors. Restriction enzyme sites are italicized. Abbreviations: *B*, *Bam*HI; *E*, *Eco*RI; *H*, *Hin*dIII; *K*, *Kpn*I; *N*, *Nco*I; *S*, *Sac*I; *X*, *Xho*I.

Table 2
List of preferred codons for *P. pastoris*

| Amino acid | Preferred codon |
|---|---|
| Alanine | GCT |
| Arginine | AGA |
| Aparagine | AAC |
| Aspartic acid | GAC, GAT |
| Cysteine | TGT |
| Glutamic acid | GAG, GAA |
| Glutamine | GAA, GAG |
| Glycine | GGT |
| Histidine | CAC |
| Isoleucine | ATT, ATC |
| Leucine | TTG |
| Lysine | AAG |
| Methionine | ATG |
| Phenylalanine | TTC |
| Proline | CCA, CCT |
| Serine | TCT, TCC |
| Threonine | ACT, ACC |
| Tryptophan | TGG |
| Tyrosine | TAC |
| Valine | GTT, GTC |
| Stop | TAA |

A preferred codon for *P. pastoris* was defined as a codon that had more than a 30% frequency for encoding an amino acid in highly expressed proteins in *P. pastoris* based on the data of Sreekrishna [19].

letters indicate restriction sites and the DNA sequences for annealing two primers, respectively. The resulting DNA fragment was digested with *Sfu*I and *Eco*RI and subcloned into the same sites in pPICZα. The resulting plasmid was designated as pPICZkt.

For making the pPICZpre vector containing only the pre signal sequence of alpha mating factor, 2 primers were designed. The 5′ primer annealing to the 5′ end of 5′ AOX1 in the pPICZα vector was GGT CAT GAG ATC AGA TCT AAC ATC. The 3′ primer, annealing to the 3′ end of the DNA sequence coding for the pre signal peptide of alpha mating factor, was GCT GAA TTC TCT TTT CTC GAG GCT GCG AAT AAA ACA GCA G. This primer created two restriction sites, *Xho*I and *Eco*RI sites (underlined letters). The *Xho*I site provided the cloning site for the immunotoxin gene and the *Eco*RI site was used for subcloning the PCR product into a pPICZ vector to obtain the pPICZpre vector.

To subclone the rebuilt gene of immunotoxin into pPIC9K, the *Sac*I–*Eco*RI fragment having a part of the 5′ *AOX1* promoter region, the prepro signal sequence of alpha mating factor, and the rebuilt gene of immunotoxin was obtained by digesting the immunotoxin construct in pnPICZα with *Sac*I and *Eco*RI. The gel purified *Sac*I–*Eco*RI fragment was inserted into the *Sac*I and *Eco*RI sites in pPIC9K.

For cytosolic expression of DT A chain, the A chain gene was inserted into a pPIC3 vector. The A chain gene was amplified by using the third version of the synthetic gene as the template DNA and then digested with *Bam*HI and *Not*I. The resulting DNA fragment was ligated into a pPIC3 vector that was treated with *Bam*HI, *Not*I, and calf intestinal phosphatase.

### *Pichia pastoris transformation and selection of transformants expressing immunotoxin*

Ten μg of plasmid encoding the bivalent anti-human anti-T-cell immunotoxin was linearized with *Sac*I and then electroporated into X-33, GS115 or KM71 strains utilizing 200 Ω, 7500 V/cm, and 25 μF with a Bio-Rad GenePulser (Bio-Rad Laboratories). Transformants expressing selectable markers were obtained by spreading onto the appropriate agar selection plate. The resulting transformants were scored for the phenotype of methanol utilization (Table 3). Twelve single colonies from each transformation were grown in 5 ml of YPD medium in 14 ml test tubes for 2 days to obtain a saturated cell density and then resuspended in 3 ml of BMMYC medium for induction. Fifteen μl of methanol was supplemented every 24 h after initiation of methanol induction. The supernatants were harvested after 2 days of methanol induction and then subjected to SDS–PAGE or Western blotting to check the immunotoxin expression level.

### *Expression of immunotoxin in shake flask culture*

For small scale shake flask culture, a single colony or frozen stock was grown in 50 ml of YPD medium in a 250 ml Erlenmeyer flask for 2 days at 30 °C with vigorous shaking, resuspended in 30 ml of BMMYC medium, and induced for 1–4 days with the addition of 150 μl of methanol every 24 h.

For large-scale culture, 50 ml of YPD medium was inoculated with 1 ml of frozen stock and grown for two days to establish a seed culture. Fifteen ml of seed culture was used to inoculate 300 ml of YPD medium in a 1000 ml Erlenmeyer flask and the inoculated medium was cultivated for 2 days. The cell pellet was resuspended in 200 ml of BMMYC medium and induced for 4 days by adding 1 ml of methanol every 24 h.

### *Purification of immunotoxin*

A two-step purification was employed to purify recombinant immunotoxin. Anion exchange chromatography was used as the first step. Culture supernatant was dialyzed against 20 mM Tris–Cl, pH 8.0, before applying 50 column volumes (CV) to a 1 ml of column of DEAE Sepharose (Pharmacia) pre-equilibrated in the same buffer. The bound immunotoxin was eluted with 1–2 CV of elution buffer (2 M NaCl and 20 mM Tris–Cl, pH 8.0). Protein L plus affinity chromatography (Pierce Chemical) was used as a polishing step employing a 3 ml column. Before loading, the sample was made 0.6 M in $(NH_4)_2SO_4$ by the addition of 3 M $(NH_4)_2SO_4$.

Table 3
Expression strains and plasmids used to express recombinant immunotoxin in *Pichia pastoris*

| Name of strain | Host | Transforming plasmids and synthetic version of gene[a] | Signal peptide | Phenotype | Transformation | Theoretical copy number | Expression level (μg/ml) |
|---|---|---|---|---|---|---|---|
| pJHW#1 | GS115 | A-dmDT390-bisFv($G_4S$) in pnPICZα[5]<br>A-dmDT390-bisFv($G_4S$) in pPIC9K[5] | α-Prepro | $Mut^+His^+Zeocin^R$ | Double | 2 | 10 |
| pJHW#2 | GS115 | A-dmDT390-bisFv($G_4S$) in pPIC9K[5] | α-Prepro | $Mut^+His^+Zeocin^S$ | Single | 1 | 10 |
| pJHW#3 | KM71 | A-dmDT390-bisFv($G_4S$) in pPIC9K[5] | α-Prepro | $Mut^SHis^+Zeocin^S$ | Single | 1 | 10 |
| pJHW#4 | X-33 | A-dmDT390-bisFv($G_4S$) in pnPICZα[5] | α-Prepro | $Mut^+His^+Zeocin^R$ | Single | 1 | 10 |
| pJHW-α | KM71 | A-dmDT390-bisFv($G_4S$) in pnPICZα[3] | α-Prepro | $Mut^SHis^-Zeocin^R$ | Single | 1 | 1–2 |
| pJHW-kt | KM71 | A-dmDT390-bisFv($G_4S$) in pPICZ-kt[3] | Killer toxin | $Mut^SHis^-Zeocin^R$ | Single | 1 | 1–2 |
| pJHW-pre | KM71 | A-dmDT390-bisFv($G_4S$) in pPICZ-pre[3] | α-Pre | $Mut^SHis^-Zeocin^R$ | Single | 1 | <0.01 |
| pJHW-EAEAEF | KM71 | EAEAEF-dmDT390-bisFv($G_4S$) in pPICZα[3] | α-Prepro | $Mut^SHis^-Zeocin^R$ | Single | 1 | 1–2 |
| pJHW-w/o | KM71 | dmDT390-bisFv($G_4S$) in pnPICZα[3] | α-Prepro | $Mut^SHis^-Zeocin^R$ | Single | 1 | 1–2 |

Letters preceding dmDT390 represent extra $NH_2$-terminal amino acids not present in the wild type toxin.
[a] Transforming plasmids are named according to gene structure followed by the parental vector. Superscript 5 indicates the final version of the synthetic gene encoding the immunotoxin. Superscript 3 indicates the third version of synthetic gene encoding the immunotoxin.

The column was washed with 1.5 CV of binding buffer (600 mM $(NH_4)_2SO_4$, 20 mM Tris–Cl, pH 8.0) and the bound immunotoxin was eluted with 5 CV of 20 mM Tris–Cl buffer (pH 8.0).

### *SDS–PAGE and Western blotting*

Proteins in culture supernatants were subjected to SDS–PAGE utilizing Tris-glycine 10% precast gels (Invitrogen) under non-reducing or reducing conditions. For Western blotting, the fractionated proteins were transferred onto nitrocellulose membranes by electroblotting. Non-specific binding was blocked with 5% nonfat milk in TBST buffer (50 mM Tris–HCl, pH 7.5, 150 mM NaCl, and 0.1% Tween 20). Goat polyclonal antibody directed against diphtheria toxin [6] diluted 1:2000 was used as the primary antibody and alkaline phosphatase-conjugated rabbit anti-goat IgG (Roche Molecular Biochemicals) diluted 1:5000 was used as the secondary antibody. The immunotoxin was visualized with one-step NBT/BCIP substrate (Pierce Chemical).

### *Protein quantification by UV absorbance at 276 nm or comparison on SDS–PAGE*

The recombinant immunotoxin was purified by a two-step purification as mentioned above and then dialyzed against a large quantity of storage buffer (200 mM NaCl, 1 mM EDTA, 5% glycerol and 20 mM Tris–Cl, pH 8.0). After filter sterilization, the concentration of immunotoxin was quantified by $OD_{276\,nm}$ using $A_{0.1\%/cm} = 1.63$, which was calculated with the Web-based program (www.up.univ-mrs.fr/~wabim/d_abim/compo-p.html) by the deduced amino acid sequence of the immunotoxin. The protein concentration of the preparation by UV absorbance was 96% of that measured by BCA Protein Assay Kit (Pierce Chemical) in which bovine serum albumin was used as a standard. The expression level of immunotoxins obtained from single colonies was quantified by SDS–PAGE using recombinant immunotoxin standards of known concentration based on UV absorbance.

### *Cytotoxicity assay*

The specific cytotoxicity of anti-human anti-CD3 immunotoxins expressed in *P. pastoris* was performed as described [3]. Briefly, immunotoxins were applied to Jurkat cells, a human $CD3\varepsilon^+$ T-cell leukemia line, ($5 \times 10^4$ cells/well) in 96-well plates in leucine-free RPMI 1640 medium. After 20 h, a 1 h pulse of $[^3H]$leucine was given. Cells were collected onto filters with a Skatron harvester. After addition of scintillant, samples were counted in a Beckman scintillation counter using standard LSC techniques.

## Results

### *DNA rebuilding for abolishing AT-rich regions and codon optimization*

The original gene of A-dmDT390-bisFv($G_4S$) placed under the *AOX1* promoter and preceded by the alpha-mating factor signal peptide was not expressed in *P. pastoris*. To find out why A-dmDT390-bisFv($G_4S$) was not expressed in *P. pastoris*, RT-PCR was carried out. The result of RT-PCR showed that a full size transcript of A-dmDT390-bisFv($G_4S$) was not transcribed. Many reports have demonstrated that an AT-rich region in a gene of interest could induce early termination of transcription in yeast including *P. pastoris*. It was found that the DNA of the DT390 moiety had three major AT-rich regions. DNA rebuilding for abolishing AT-rich regions was started from the 3′ end, because Romanos and his colleagues [13] pointed out that known transcription terminators present near the 5′ end of an open reading frame did not efficiently induce early termination of transcription.

The most important constructs that we expressed in the course of optimizing immunotoxin expression are listed in Table 3 along with the host strain, phenotype, synthetic gene version number, signal peptide, expression level, and the number of transformations. Several amino-terminal variants are also included. As described in more detail below, the synthetic gene version was a major variable in increasing the expression from undetectable levels to 10 μg/ml.

The toxin domains, A-dmDT390, contain three major AT-rich regions of 65.6%, 69.2%, and 63.8% from the 5′ to the 3′ position, respectively (see Fig. 1, top). The 3′ region was rebuilt first, then the middle region was added, and the third rebuilding added the 5′ region (see Materials and methods for details). Fig. 1 lists the final AT content of each region. Expression was not achieved until the third rebuilding and was in the range of 1–2 μg/ml (Table 3). Rebuilding was then performed on two minor AT-rich regions in the toxin domains (see example 4 in Fig. 1). The AT content in these regions was reduced from 60.2% to 51.4% in the A chain and 66.7–53.8% in the truncated B chain. In the course of this rebuilding, codon optimization was performed on the adjacent regions (see Table 2). However, the expression level remained at 1–2 μg/ml. The fifth and final rebuilding reduced an AT-rich region in the variable light chains of both sFvs from 67.9% to 56.8% and a similar region in the variable heavy chains from 64.6% to 56.3%. Adjacent codons were optimized and rare *P. pastoris* codons in the linker regions were optimized (see example 5 of Fig. 1). The fifth rebuilding changes, when combined with the previous changes 1–4, resulted in a 5-fold increase in production over the third and fourth rebuilding (Table 3).

### *The relative tolerance of P. pastoris to DT*

The tolerance of *P. pastoris* to DT was demonstrated by expressing the DT catalytic A chain within the cytosolic compartment of *P. pastoris* and measuring colony counts and A chain expression after induction. The A chain gene of DT was placed into the pPIC3 vector under control of the *AOX1* promoter. This vector lacks a signal sequence and is used for cytosolic expression in *P. pastoris*. Among the transformants, 12 colonies were randomly selected for screening for A chain production and all were A chain positive by Western blotting. Strain C-4 showed the highest level of A chain production.

To measure colony forming units, we used the wild type X-33 strain as a control, pJHW#2 strain for expressing bivalent immunotoxin and C-4 strain expressing A chain. These three strains were cultivated in 50 ml of YPD for 2 days at 30 °C and the harvested cell pellets were resuspended in methanol induction media to give 8% wet cell density. For expression of bivalent immunotoxin or A chain, each culture was induced for 24 h in the presence of methanol. After 24 h, samples from each culture were diluted $10^4$–$10^6$-fold with PBS buffer and 100 μl of aliquot was plated onto a YPD plate. Following 3 days of incubation at 30 °C, the number of colonies appearing on YPD plates was counted and the relative colony forming units (%) of each strain to the wild type strain was calculated. Among the colonies recovered after 24 h of methanol induction, 12 colonies were randomly checked to determine whether they still had the ability to express bivalent immunotoxin or A chain of DT by Western blotting. A chain expression was assayed from lysed yeast cells while bivalent immunotoxin was assayed from the culture supernatant. All recovered colonies still had the ability to express bivalent immunotoxin or A chain of DT although colony forming units were significantly lowered as compared with the wild type X-33 strain (see Fig. 3). Loss of colony forming units was approximately 55% for bivalent immunotoxin expression via the secretory route and 95% for cytosolic A chain expression. The expression of A chain in these experiments was about 1/300 of the expression of bivalent immunotoxin per ml of culture or about 33 ng/ml. The growth of *P. pastoris* and the secretion of recombinant immunotoxin in *P. pastoris* in shake flask culture were not affected by adding DT to the medium at a concentration of 300 μg/ml.

In a previous study of DT based immunotoxin expression in *S. cerevisiae*, expression of cytosolic DT A chain under the *GAL1* promoter was not observed following induction unless the strain had been previously mutated to toxin resistance [20]. And, the survival of wild type *S. cerevisiae* cells expressing the DT A chain gene was $1.1 \times 10^{-6}$ cells per generation. Thus, *P. pastoris* appears to be relatively tolerant to DT compared to *S. cerevisiae* and wild type CHO cells. Tolerance to DT

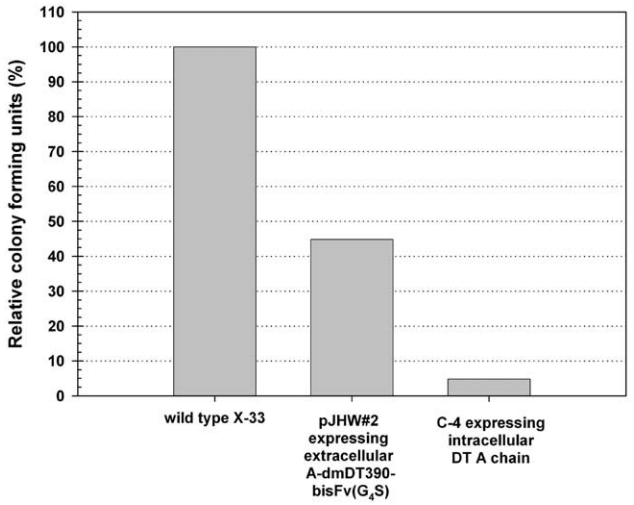


Fig. 3. Relative colony forming units (%) were decreased by expression of toxin either in the cytosol or by the secretory route. Each value was obtained after 24 h of methanol induction.

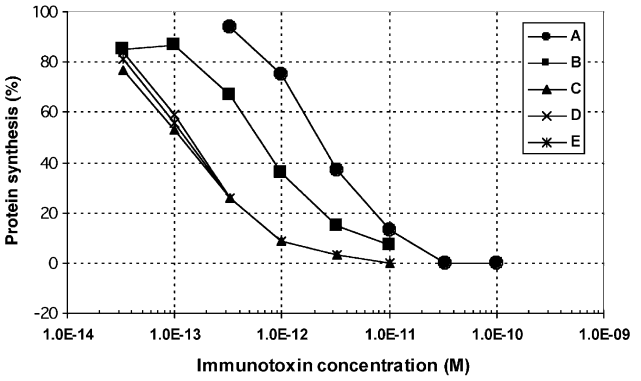


Fig. 4. Comparison of specific toxicity towards Jurkat T cells between immunotoxin constructs having a different number of amino acid residues at $NH_2$-terminus of the DT A chain. A, UCHT1-CRM9 chemical conjugate; B, EAEAEP-dmDT390-bisFv($G_4S$) from *P. pastoris*; C, A-dmDT390-bisFv($G_4S$) from *P. pastoris*; D, dmDT390-bisFv($G_4S$) from *P. pastoris*; E, A-dmDT390-bisFv($G_4S$) from CHO cells.

appears to be enhanced by utilizing the secretory pathway of expression.

*Effect of additional $NH_2$-terminal amino acids on immunotoxin potency*

The bivalent anti-T-cell immunotoxin, A-dmDT390-bisFv($G_4S$), was first expressed in a mutated CHO cell line [5]. In this experiment, the secretional signal peptide of mouse κ-immunoglobulin was used for secretion of the immunotoxin. For optimal cleavage of signal peptide by signal peptidase, an additional amino acid, alanine, was added at the N-terminus of the immunotoxin. The first, second, and third versions of the synthetic gene were subcloned into the *Eco*RI site of pPICZα, resulting in the inclusion of unnecessary amino acid residues (Glu–Phe, Glu–Ala–Glu–Phe or Glu–Ala–Glu–Ala–Glu–Phe) at $NH_2$-terminus of immunotoxin depending on the degree cleavage of the EAEA linker by STE13 peptidase. We also constructed a gene for dmDT390-bisFv($G_4S$) that did not code for an additional amino acid residue at the $NH_2$-terminus, to compare the toxicity of these constructs. Previous work had shown that the translocation of DT A chain to the cytosol could be attenuated by the addition of extra amino acid residues at toxin $NH_2$-terminus [22]. To compare the efficacy of these three constructs, toxicity assays were employed on purified immunotoxins obtained from the supernatants of each individual transformant as described in Materials and methods. Fig. 4 indicates that an alanine $NH_2$-residue does not affect toxicity, but the acidic 6 residue $NH_2$-sequence decreased toxicity. In addition, we noted that this construct when purified by anion exchange chromatography at pH 6.4 as opposed to pH 7.4 lost 90% of its bioactivity.

*Effect of different signal peptides on secretion level of immunotoxin*

The type of signal peptide used for secretion of the protein of interest is a variable, which can improve expression level. In this study, the prepro secretion peptide of alpha-mating factor was used for secretion of immunotoxin. We also compared the secretion peptide of killer toxin from *Kluyveromyces lactis* and the pre secretion peptide of alpha-mating factor which can secrete interleukin 1β [21] and aminoglycoside phosphotransferase (APH), human granulocyte-macrophage colony stimulating factor (hGH-CSF), and interleukin β in S. cerevisiae [23], respectively. Alpha prepro mating factor and killer toxin were equally effective in secreting immunotoxin while the α-pre sequence was not effective (Table 3).

*Optimization of composition of induction media for production of immunotoxin*

The immunotoxin was expressed only in a complex medium such as BMMY medium or BMMY medium plus casamino acids as described in Materials and methods, but not in a defined medium. DT based immunotoxins such as A-dmDT390-bisFv($G_4S$) are very sensitive to low pH because exposure of internal hydrophobic helicies occurs in a proton-dependent manner [24,25]. We found that a pH above 6.0 must be maintained to inhibit irreversible conformational change of the DT390 moiety. At this pH, some proteases are likely

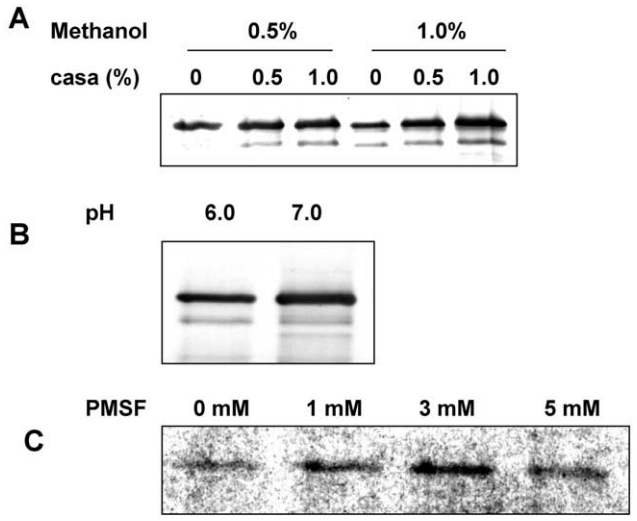


Fig. 5. Effects of media pH, casamino acids, methanol concentration, and PMSF on expression of anti-human anti-T-cell immunotoxin, A-dmDT390-bisFv($G_4S$). Panels A and B were obtained from Western blots and panel C was from a Coomassie blue-stained SDS–PAGE.

to be present in active forms, derived either from secretion or breakdown of *P. pastoris* cells. Compared to defined medium, complex media have abundant amino acids and peptides that can inhibit proteases. Casamino acids have been previously used to inhibit protease activity during *P. pastoris* fermentation [19]. With the addition of casamino acids in the methanol induction medium, the expression level of immunotoxin was improved as shown in Fig. 5. In addition, the expression level was increased 2-fold on raising the pH from 6.0 to 7.0 We also tested PMSF and EDTA as protease inhibitors. PMSF was effective at 1–3 mM (Fig. 5) whereas EDTA was not effective (data not shown).

*Effect of gene copy number on expression level*

To determine the effect of copy number of the immunotoxin gene on the expression level, double copy transformants were obtained by double transformation of the immunotoxin constructs in pnPICZα and pPIC9K. There was no difference in expression level between single and double copy transformants (Table 3). Methanol utilization in the double copy transformants was unchanged compared to single copy transformants.

*Protein purification*

Two-step chromatography was employed to purify A-dmDT390-bisFv($G_4S$). DEAE Sepharose ion exchange chromatography was used for the concentration of immunotoxin. Protein L affinity chromatography in the presence of high salt was used as a polishing step. Although the immunotoxin bound weakly under these conditions, substantial purification was obtained and the immunotoxin was eluted in low salt. The eluted material (Fig. 6) ran as a single band of the expected molecular weight (96.3 kDa) in a non-reducing SDS gel. The purified immunotoxin was fully functional when assayed by its specific toxicity in a protein synthesis assay in Jurkat cells (Fig. 4).

Because the binding of immunotoxin to Protein L was weak, it was necessary to limit the loading volume to less than 1 CV of Protein L resin and the washing volume to 1.5 CV. Many studies have reported that sFvs can be purified by Protein L agarose affinity chromatography. The binding affinity of Protein L to immunoglobulin light chains is predominantly directed at specific subtypes of kappa chains [26,27]. A comparative study of amino acid sequences between the binding group and the non-binding group of selectively labeled $V_L$-domains of murine Fab fragments showed that the residues at positions 9, 20, and 74 of the light chain variable region were important in binding to Protein L [28]. Within the light chain variable regions of sFv(UCHT1) which is a mouse κV subtype, these critical amino acid residues responsible for Protein L bind-

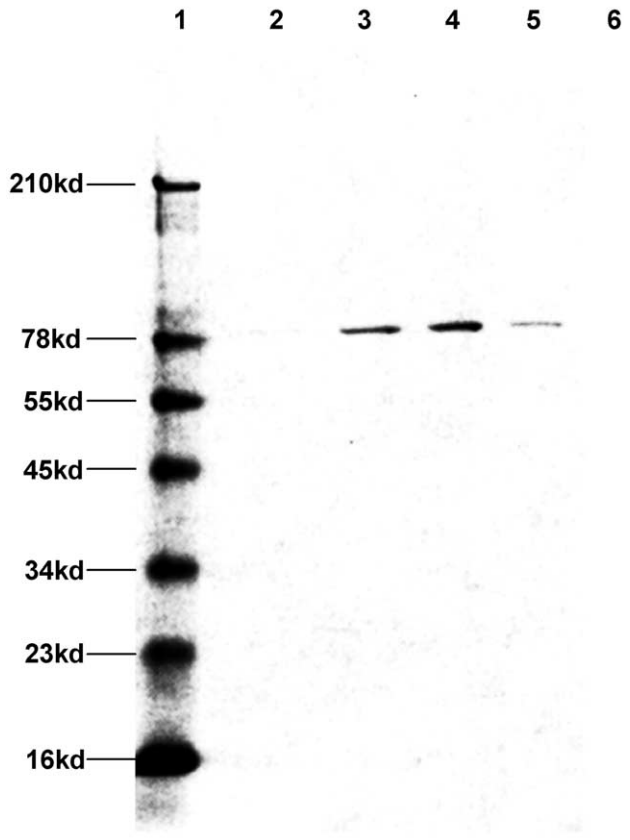


Fig. 6. SDS–PAGE of fractions obtained from Protein L affinity chromatography. Lane 1, See Blue (Invitrogen) pre-stained protein markers, lanes 2–6, fraction #4 to #8 from Protein L affinity chromatography. Fractions #5 and #6 were pooled and then used for the protein synthesis assay of Fig. 4.

ing were identical to those of the group binding Protein A. However, A-dmDT390-bisFv($G_4S$) bound weakly to Protein L and only under high salt conditions. The binding and elution patterns of the immunotoxin in Protein L affinity chromatography were very similar to those of hydrophobic interaction chromatography. An interaction may occur between the hydrophobic patch of sFv in the immunotoxin and Protein L. This hydrophobic interaction may be facilitated under high salt conditions and may also involve the DT390 moiety.

## Discussion

Most diphtheria toxin or pseudomonas exotoxin-based fusion proteins have been expressed in *E. coli* and refolded in vitro from solubilized cytoplasmic inclusion bodies. This system was used because these toxins are not active against prokaryotic protein synthetic machinery and the relatively simple domain structure of monovalent immunotoxins is amenable to in vitro refolding. However, the optimized bivalent anti-human anti-T-cell immunotoxin having a multi-domain structure was not efficiently refolded at full activity from *E. coli*. In a previous study, we expressed monovalent and bivalent anti-human anti-T-cell immunotoxins in the mutated CHO cell line CHO K1 RE1.22c [5,6] utilizing a kappa chain secretion signal. However, recent attempts to optimize this CHO expression revealed a feedback inhibition limitation of expression that suggested retrograde transport of the secreted product with subsequent inhibition of secretion that occurs in mutated CHO cells.

We employed *P. pastoris* as an alternative expression system to produce bivalent anti-human anti-T-cell immunotoxin. *P. pastoris* has been successfully used for the expression of various heterologous proteins used in therapeutics such as vaccines [13,14,29,30], single chain antibodies [31,32], and human-derived peptides or protein [33–35]. It is known that the A chain (catalytic domain) of diphtheria toxin is not measurably expressed and is toxic in most eukaryotes including *S. cerevisiae* due to the catalytic inactivation of EF-2. In contrast to *S. cerevisiae* [20], A chain of diphtheria toxin was expressed in the cytosol of wild type *P. pastoris,* and colony forming units were detected after induction, although A chain-expressing clones were decreased by 95% after 24 h. Yet, all surviving cells had the ability to express A chain in the cytosol. These results suggest that *P. pastoris* has a certain level of tolerance to DT not seen in wild type vertebrate cells or *S. cerevisiae*. The mechanism of this tolerance is not known. It could reflect differences in the relative rates of production and degradation of A chain and of EF-2 between *P. pastoris* on the one hand and *S. cerevisiae* and CHO cells on the other. The fact that the secretion of A-dmDT390-bisFv($G_4S$) by *P. pastoris* only reduced colony forming units by 55% after 24 h indicates that tolerance to DT is enhanced via the secretory pathway. This could reflect two processes. First, a very tight coupling between translation and secretion keeping the cytosol compartment free of toxin, and second, the relative absence of retrograde transport after secretion aided by the physical barrier of the yeast polysaccharide capsule. We observed that like intact *S. cerevisiae, P. pastoris* growth and secretion of recombinant immunotoxin are not effected by high doses of external DT. The fact that yeast spheroplasts are sensitive to high concentrations of DT [11,36] suggests that the yeast cell wall plays a role in limiting retrograde transport. This may be one explanation for the higher expression levels of A-dmDT390-bisFv($G_4S$) in *P. pastoris*, up to 35 mg/l in fermentor production, compared to toxin resistant CHO cells (Woo et al., manuscript in preparation).

In this report, we discovered that expression of the bivalent anti-human anti-T-cell immunotoxin (A-dmDT390-bisFv($G_4S$)) in *P. pastoris* could be improved by gene optimization for abolishing AT-rich regions in the original DNA sequence of immunotoxin and codon optimization for efficient translation. The expression of the original DNA sequence of immunotoxin failed in *P. pastoris* due to the early termination of transcription. The DNA sequence analysis shows that there were two highly conserved AATAAA hexanucleotide sequences at 172nd and 682nd position of the DNA sequence of the DT390 moiety. In higher eukaryotic cells, this conserved nucleotide sequence is located ~20 nucleotides upstream of the site of cleavage and poly(A) addition. The second component of the polyadenylation signal (a run of Ts or a poorly defined GT-rich sequence) is located just downstream of the poly(A) site [37,38]. Although the second component for termination of transcription was not found downstream of AATAAA consensus sequence in the immunotoxin gene, three major AT-rich regions present in the DNA sequence of DT390 could induce the early termination of transcription in *P. pastoris*. The issue of early termination of transcription in *P. pastoris* and *S. cerevisiae* has been previously overcome by lowering the AT content in AT-rich region by DNA rebuilding [13,14]. In the third version of DNA rebuilding work, the AT content of DNA sequences of the three AT-rich regions was reduced from more than 64% to less than 50% and simultaneously codon usage was changed to a preferred codon usage for *P. pastoris*. In the fifth and final rebuilding, DNA sequences in minor AT-rich regions having 50–90 nucleotides and non-preferred codon rich regions such as $(G_4S)_3$ artificial linkers were replaced by the DNA sequences of preferred codons. Table 2 shows a list of preferred codons for *P. pastoris*. Sreekrishna [19] has analyzed the frequency of codon usage in the highly expressed *P. pastoris* genes *AOX*1, *AOX*2, dihydroxy acetone synthetases, 1 and 2, and glyceralde-

hyde phosphate dihydrogenase genes. Based on these data, a preferred codon was defined as a codon that had more than a 30% frequency for encoding an amino acid in the highly expressed proteins in *P. pastoris*. Subsequently, gene optimization increased expression level up to 10 mg/L in shake flask culture. However, in this fifth and final rebuilding version as in previous versions, we did not determine the relative contributions of codon optimization versus AT content reduction in the 5-fold enhanced protein expression. Therefore, we cannot draw any conclusion as to the utility of codon optimization for enhancing production beyond that achieved by % AT reduction.

Since the DT390 moiety of our immunotoxin undergoes an irreversible conformational change at pH values below 6.0, it was expressed at pH 7.0 in *P. pastoris*. The expression level was improved by adding the protease inhibitor PMSF and by adding casamino acids. On the other hand, the expression level was not improved by using proteinase-deficient strain, SMD1168 (*his*4 *pep*4).

In conclusion, our results demonstrate that *P. pastoris* is an ideal expression system for expression of a toxin-based fusion protein and gene optimization is an efficient way to improve the expression level.

## References


[1] J.M. Thomas, D.E. Eckhoff, J.L. Contreras, A.L. Lobashevsky, W.J. Hubbard, J.K. Moore, W.J. Cook, F.T. Thomas, D.M. Neville Jr., Durable donor-specific T and B cell tolerance in rhesus macaques induced with peritransplantation anti-CD3 immunotoxin deoxyspergualin: absence of chronic allograft nephropathy, Transplantation 69 (2000) 2497–2503.

[2] J.M. Thomas, J.L. Contreras, C.A. Smyth, A. Lobashevsky, S. Jenkins, W.J. Hubbard, D.E. Eckhoff, S. Stavrou, D.M. Neville Jr., F.T. Thomas, Successful reversal of streptozotocin-induced diabetes with stable allogeneic islet function in a preclinical model of type 1 diabetes, Diabetes 50 (2001) 1227–1236.

[3] D.M. Neville Jr., J. Scharff, K. Srinivasachar, In vivo T-cell ablation by a holo-immunotoxin directed at human CD3, Proc. Natl. Acad. Sci. USA 89 (1992) 2585–2589.

[4] D.M. Neville Jr., J. Scharff, H.Z. Hu, K. Rigaut, J. Shiloach, W. Slingerland, M. Jonker, A new reagent for the induction of T-cell depletion, anti-CD3-CRM9, J. Immunother. Emphasis Tumor Immunol. 19 (1996) 85–92.

[5] J. Thompson, S. Stavrou, M. Weetall, J.M. Hexham, M.E. Digan, Z. Wang, J.H. Woo, Y. Yu, A. Mathias, Y.Y. Liu, S. Ma, I. Gordienko, D.M. Neville Jr., Improved binding of a bivalent single-chain immuntoxin results in increased efficacy for in vivo T-cell depletion, Protein Eng. 14 (2001) 1035–1041.

[6] J. Thompson, H. Hu, J. Scharff, D.M. Neville Jr., An anti-CD3 single-chain immunotoxin with a truncated diphtheria toxin avoids inhibition by pre-existing antibodies in human blood, J. Biol. Chem. 270 (1995) 28037–28041.

[7] J.M. Hexham, D. Dudas, R. Hugo, J. Thompson, V. King, C. Dowling, D.M. Neville Jr., M.E. Digan, P. Lake, Influence of relative binding affinity on efficacy in a panel of anti-CD3 scFv immunotoxins, Mol. Immunol. 38 (2001) 397–408.

[8] Y.Y. Liu, I. Gordienko, A. Mathias, S. Ma, J. Thompson, J.H. Woo, D.M. Neville Jr., Expression of an anti-CD3 single-chain immunotoxin with a truncated diphtheria toxin in a mutant CHO cell line, Protein Expr. Purif. 19 (2000) 304–311.

[9] M. Yamaizumi, E. Mekada, T. Uchida, Y. Okada, One molecule of diphtheria toxin fragment A introduced into a cell can kill the cell, Cell 15 (1978) 245–250.

[10] D. O'Keefe, R.J. Collier, Cloned diphtheria toxin within the periplasm of *Escherichia coli* causes lethal membrane damage at low pH, Proc. Natl. Acad. Sci. USA 86 (1989) 343–346.

[11] S. Murakami, J.W. Bodley, D.M. Livingston, *Saccharomyces cerevisiae* spheroplasts are sensitive to the action of diphtheria toxin, Mol. Cell. Biol. 2 (1982) 588–592.

[12] G. Gellissen, Heterologous protein production in methylotrophic yeasts, Appl. Microbiol. Biotechnol. 54 (2000) 741–750.

[13] M.A. Romanos, A.J. Makoff, N.F. Fairweather, K.M. Beesley, D.E. Slater, F.B. Rayment, M.M. Payne, J.J. Clare, Expression of tetanus toxin fragment C in yeast: gene synthesis is required to eliminate fortuitous polyadenylation sites in AT-rich DNA, Nucleic Acids Res. 19 (1991) 1461–1467.

[14] M.P. Byrne, R.W. Titball, J. Holley, L.A. Smith, Fermentation, purification, and efficacy of a recombinant vaccine candidate against botulinum neurotoxin type F from *Pichia pastoris*, Protein Expr. Purif. 18 (2000) 327–337.

[15] M.L. Valencik, J.A. McDonald, Codon optimization markedly improves doxycycline regulated gene expression in the mouse heart, Transgenic Res. 10 (2001) 269–275.

[16] M. Massaer, P. Mazzu, M. Haumont, M. Magi, V. Daminet, A. Bollen, A. Jacquet, High-level expression in mammalian cells of recombinant house dust mite allergen ProDer p 1 with optimized codon usage, Int. Arch. Allergy Immunol. 125 (2001) 32–43.

[17] D.P. Humphreys, M. Sehdev, A.P. Chapman, R. Ganesh, B.J. Smith, L.M. King, D.J. Glover, D.G. Reeks, P.E. Stephens, High-level periplasmic expression in *Escherichia coli* using a eukaryotic signal peptide: importance of codon usage at the 5′ end of the coding sequence, Protein Expr. Purif. 20 (2000) 252–264.

[18] V.S. Te'o, A.E. Cziferszky, P.L. Bergquist, K.M. Nevalainen, Codon optimization of xylanase gene xynB from the thermophilic bacterium *Dictyoglomus thermophilum* for expression in the filamentous fungus *Trichoderma reesei*, FEMS Microbiol. Lett. 190 (2000) 13–19.

[19] K. Sreekrishna, Strategies for optimizing protein expression and secretion in the methylotrophic yeast *Pichia pastoris*, in: R.H. Baltz, G.D. Hegeman, P.L. Skatrud (Eds.), Industrial Microorganism: Basic and Applied Molecular Genetics, Am. Soc. Microbiol., Washington, DC, 1993, pp. 119–126.

[20] J.P. Perentesis, F.S. Genbauffe, S.A. Veldman, C.L. Galeotti, D.M. Livingston, J.W. Bodley, J.R. Murphy, Expression of diphtheria toxin fragment A and hormone-toxin fusion proteins in toxin-resistant yeast mutants, Proc. Natl. Acad. Sci. USA 85 (1988) 8386–8390.

[21] C. Baldari, J.A. Murray, P. Ghiara, G. Cesareni, C.L. Galeotti, A novel leader peptide which allows efficient secretion of a fragment of human interleukin 1 beta in *Saccharomyces cerevisiae*, EMBO J. 6 (1987) 229–234.

[22] I.H. Madshus, S. Olsnes, H. Stenmark, Membrane translocation of diphtheria toxin carrying passenger protein domains, Infect. Immun. 60 (1992) 3296–3302.

[23] J.F. Ernst, Efficient secretion and processing of heterologous proteins in *Saccharomyces cerevisiae* is mediated solely by the presegment of alpha-factor precursor, DNA 7 (1988) 355–360.

[24] K. Sandvig, S. Olsnes, Diphtheria toxin entry into cells is facilitated by low pH, J. Cell Biol. 87 (1980) 828–832.

[25] R.K. Draper, M.I. Simon, The entry of diphtheria toxin into the mammalian cell cytoplasm: evidence for lysosomal involvement, J. Cell Biol. 87 (1980) 849–854.

[26] B.H. Nilson, A. Solomon, L. Bjorck, B. Akerstrom, Protein L from *Peptostreptococcus magnus* binds to the kappa light chain variable domain, J. Biol. Chem. 267 (1992) 2234–2239.

[27] B. Akerstrom, B.H. Nilson, H.R. Hoogenboom, L. Bjorck, On the interaction between single chain Fv antibodies and bacterial immunoglobulin-binding proteins, J. Immunol. Methods 177 (1994) 151–163.

[28] J. Enokizono, M. Wikstrom, U. Sjobring, L. Bjorck, S. Forsen, Y. Arata, K. Kato, I. Shimada, NMR analysis of the interaction between protein L and Ig light chains, J. Mol. Biol. 270 (1997) 8–13.

[29] J.M. Cregg, J.F. Tschopp, C. Stillman, R. Siegel, M. Akong, W.S. Craig, R.G. Buckholz, K.R. Madden, P.A. Kellaris, G.R. Davis, B.L. Smiley, J. Cruze, R. Torregrossa, G. Velicelebi, G.P. Thill, High-level expression and efficient assembly of hepatitis B surface antigen in the methylotropic yeast, *Pichia pastoris*, Bio/Technology 5 (1987) 479–485.

[30] B.P. Peeters, O.S. de Leeuw, I. Verstegen, G. Koch, A.L. Gielkens, Generation of a recombinant chimeric Newcastle disease virus vaccine that allows serological differentiation between vaccinated and infected animals, Vaccine 19 (2001) 1616–1627.

[31] C. Marty, P. Scheidegger, K. Ballmer-Hofer, R. Klemenz, R.A. Schwendener, Production of functionalized single-chain Fv antibody fragments binding to the ED-B domain of the B-isoform of fibronectin in *Pichia pastoris*, Protein Expr. Purif. 21 (2001) 156–164.

[32] S. Hellwig, F. Emde, N.P. Raven, M. Henke, P. van Der Logt, R. Fischer, Analysis of single-chain antibody production in *Pichia pastoris* using on-line methanol control in fed-batch and mixed-feed fermentations, Biotechnol. Bioeng. 74 (2001) 344–352.

[33] M.W. Werten, W.H. Wisselink, T.J. Jansen-Van Den Bosch, E.C. de Bruin, F.A. de Wolf, Secreted production of a custom-designed, highly hydrophilic gelatin in *Pichia pastoris*, Protein Eng. 14 (2001) 447–454.

[34] Y. Wang, Z.H. Liang, Y.S. Zhang, S.Y. Yao, Y.G. Xu, Y.H. Tang, S.Q. Zhu, D.F. Cui, Y.M. Feng, Human insulin from a precursor overexpressed in the methylotrophic yeast *Pichia pastoris* and a simple procedure for purifying the expression product, Biotechnol. Bioeng. 73 (2001) 74–79.

[35] T. Boehm, S. Pirie-Shepherd, L.B. Trinh, J. Shiloach, J. Folkman, Disruption of the KEX1 gene in *Pichia pastoris* allows expression of full-length murine and human endostatin, Yeast 15 (1999) 563–572.

[36] Y. Kimata, S. Harashima, K. Kohno, Expression of non-ADP-ribosylatable, diphtheria toxin-resistant elongation factor 2 in *Saccharomyces cerevisiae*, Biochem. Biophys. Res. Commun. 191 (1993) 1145–1151.

[37] E. Wahle, W. Keller, The biochemistry of polyadenylation, Trends Biochem. Sci. 21 (1996) 247–250.

[38] C. Phillips, C.B. Kyriakopoulou, A. Virtanen, Identification of a stem-loop structure important for polyadenylation at the murine IgM secretory poly(A) site, Nucleic Acids Res. 27 (1999) 429–438.